PharmaSUG China 2026 - Paper DS-173

The Effective Preparation of NMPA Submission Package

Fu Wang, Parexel International

ABSTRACT

This paper outlines the procedural steps, roles, and responsibilities integral to preparing the National Medical Products Administration (NMPA) submission package, which is based on the translation of foreign language source materials. The document presents a comprehensive overview of quality assurance and proofreading protocols applied to vendor-translated content. Additionally, a Visual Basic for Applications (VBA) tool has been developed and implemented to enhance the efficiency of quality review processes and facilitate automated corrections for standardized dataset and variable labels where applicable. The paper details the underlying logic and functionality of this VBA tool. Furthermore, critical quality assurance considerations are discussed to ensure the integrity and completeness of the NMPA submission package.

INTRODUCTION

The market availability of pharmaceutical products and therapeutic interventions often varies across countries and regions due to multiple factors. For drugs that have received regulatory approval from the Food and Drug Administration (FDA) or European Medicines Agency (EMA), global sponsors may submit a New Drug Application (NDA) to China utilizing the same submission package previously submitted to the FDA or EMA, with or without supplementary materials. Additionally, the Center for Drug Evaluation at the NMPA released provisional guidelines on the submission of clinical trial data in July 2020.[1] From both regulatory and operational perspectives, translation of key submission package components from foreign languages into Chinese is necessary to meet NMPA requirements for multinational pharmaceutical companies.

Statistical programming teams typically lack expertise in medical translation, so translation work is often outsourced to external vendors. However, external medical translators may lack study-specific knowledge, resulting in inaccurate or suboptimal translations into Chinese. To address this gap and ensure high-quality, readable NMPA submission materials, we have established a quality review team comprising statistical programmers to conduct rigorous review and proofreading of all Chinese translations provided by external vendors.

IMPLEMENTATION

Several roles and contents mentioned in this paper are referenced below.

NMPA: National Medical Products Administration (China)

CMAT: Chinese MDDT Assistant Tool (CMAT)

CSP: China Study Programmer

NMPA Submission Package Workflow Overview

Different steps and responsibilities involved in this NMPA submission package preparation process is outlined below in ***Figure 1***. The details of each step and attention points are followed in sequence.

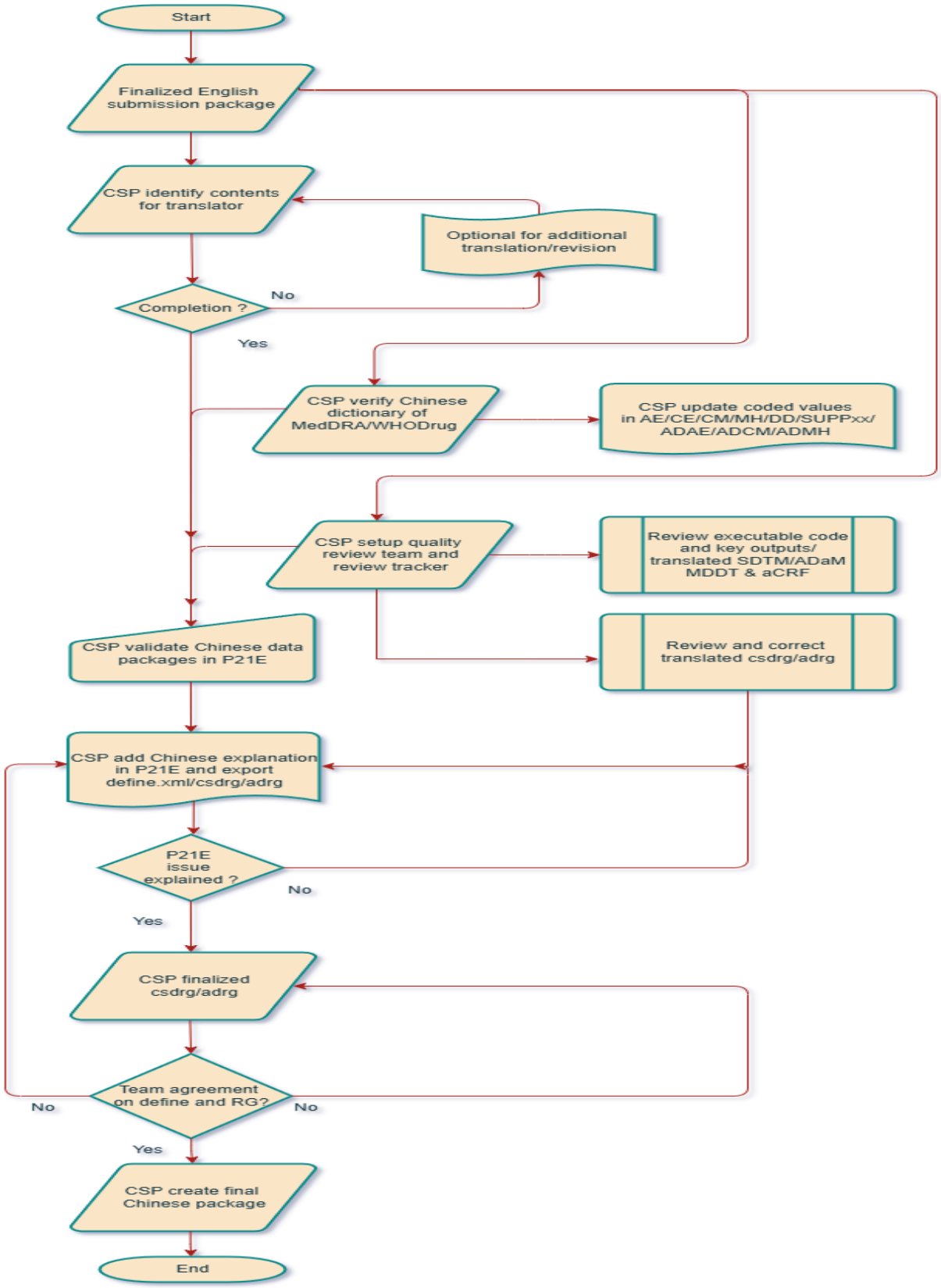

Figure 1 NMPA Submission Package Preparation Workflow

Finalization of English Package

Once the internal timeline for NMPA submission becomes available, the CSP can initiate the establishment of a quality review team and create a review tracker while the original study team finalizes the English submission package. Upon completion of the English package, the CSP should commence a comprehensive review of the executable codes and corresponding key outputs to verify that the codes are macro-free and that outputs are consistent between the macro-inclusive and macro-free versions. Since the executable macro-free codes significantly impact the macros section of review guides, it is advisable to finalize these codes concurrently with the English submission package rather than deferring them for subsequent quality review.

CSP Identify Contents And Send Out to Vendor

The CSP will identify content requiring translation in accordance with NMPA requirements. Current practice involves translating the following documents into Chinese: aCRF, CSDRG, ADRG, SDTM and ADaM EXCEL SPEC (exported from P21E/ADAE.CQzzNAM/ADAE.SMQzzNAM/ADCM.MOIzzNAM), and additional vital files such as Analysis Results Metadata (ARM), where applicable.

For aCRF translation, annotations must be added as PDF comments rather than embedded directly within the PDF file. For EXCEL SPEC exported from P21E, translators must strictly adhere to the internal meta translation file for dataset and variable labels, where available, to ensure alignment with the CDISC Chinese version, which can be subsequently verified using the CMAT tool. The EXCEL SPEC sheets for 'Datasets', 'Variables', 'ValueLevel', 'Methods' and 'Comments' are translated with consideration for both reviewer readability and sponsor workload. For the 'Datasets', 'Variables', and 'ValueLevel' sheets, English labels are translated into Chinese and recorded in a new column. Similarly, for the 'Methods' and 'Comments' sheets, English descriptions are translated into Chinese and saved in a new column.

Depending on time constraints, the CSP may or may not require the translator to revise or translate additional documents following the initial delivery of translated materials. In certain cases, the quality review team will continue to update, correct, and/or translate portions of files in response to plan modifications or identified needs.

Chinese Package Dictionaries Replacement

The CSP should verify the availability of Chinese dictionary versions (e.g., MedDRA, WHODrug) as early as possible, ideally immediately following English package finalization. If a corresponding Chinese dictionary version does not exist for the study, the submission team must reach agreement with the study team on whether to use the latest available Chinese version or the first available Chinese version.

Generally, coded terms in the following domains should be replaced with their Chinese equivalents: AE, CE, CM, DD, MH, SUPPxx, ADAE, ADCM, and ADMH, along with ADAE.SMQzzNAM, ADAE.CQzzNAM and ADAE.CQzzNAM when applicable.

For any coded terms that cannot be replaced using the available Chinese dictionary versions, the submission team must document and agree on any additional actions required.

Quality Review And Proofreading

Pre-Translation Phase

Prior to the translator's return of Chinese files, the CSP should initiate the quality review team, establish a review tracker, and allocate assignments. Additionally, the CSP should perform a comprehensive review of the executable macro-free codes and verify that their outputs are consistent with those generated by standard codes containing macros. This verification will encompass ADaM codes and key TFL codes. Various approaches may be employed for developing macro-free executable codes: either completely rewriting the ADaM and key TFL code to eliminate macros, or consolidating all macros into a single file for inclusion in individual programs. While perspectives on macro-free executable codes may vary, there is consensus that items not intended for regulatory submission shall not be referenced in SAS code. Examples include attribution macros based on MDDT specifications for assigning length and labels, and report macros that read external title and footnote files for output formatting adjustments.

Post-Translation Phase

Upon receipt of translated Chinese files from the translator, the CSP should commence quality review and proofreading. The review should prioritize EXCEL SPEC files for SDTM and ADaM, followed by aCRF, CSDRG, and ADRG. This sequencing enables the study team to validate Chinese datasets in P21E at the earliest opportunity and address any warning messages generated by the NMPA validation engine, thereby facilitating the timely export of Chinese define.xml files.

Quality Review Optimization

During the quality review and proofreading of EXCEL SPEC files for dataset and variable labels, a utility tool was developed to enhance efficiency and standardize label values across products.

The Chinese MDDT Assistant Tool (CMAT) is a VBA-based application that verifies translated dataset and variable labels against the internal meta translation file and performs automated corrections using values from the meta translation file when discrepancies are identified. Additionally, CMAT enforces the 40-byte length restriction applicable to dataset and variable labels. When a label exceeds this length limit, the tool highlights the corresponding cells to enable users to revise labels as necessary. While a live demonstration of this tool may not be available, the underlying logic and results are detailed below.

CMAT tool import files as stated below and process SDTM/ADaM EXCEL SPEC separately

SDTM EXCEL SPEC & Meta Translation File

OR

ADaM EXCEL SPEC & Meta Translation File

Actions performed by CMAT tool for sheets of ‘Datasets’/’Variables’ are described below with detailed logic, the corresponding results coming from translator without any modification and with CMAT processing are presented in ***Table 1*** and ***Table 2*** respectively for a comparison review.

- Three extra columns of 'CSP_Comments'/'Vendor_Text'/'CSP_Discussion_YN' will be added at the end within each sheet of EXCEL SPEC

- Verify labels of 'Datasets'/'Variables' Sheets against meta translation file which is aligned with CDISC Chinese version

- If datasets or variable names exist in meta file, then mark corresponding label cells background as grey, otherwise leave cell background as is

- For grey cells, if translation text doesn’t match with meta translation file, the CMAT tool replaces cell values from meta translation file with font highlighted in red and copy original translation text into column of 'Vendor_Text' and assign 'Follow Meta Translation File' in column 'CSP_Comments'

- For any translation label text in Sheets of 'Datasets'/'Variables' and in Sheets of 'ValueLevel' with dataset name like SUPPxx that is great than 13 (considering 1 Chinese character occupies 3 bytes in UTF-8 coding), CMAT tool will highlight cell background as yellow and copy original Chinese text into column of 'Vendor_Text' and assign 'Meeting length<=40 rule' into column of 'CSP_Comments', user need to manually review and shorten the Chinese texts if needed from the most conservative perspective

	A	B	C	D	E	W	X	Y
1	Order	Dataset	Variable	Label in Chinese	Label			
2	1	ADAE	STUDYID	研究标识符	Study Identifier			
3	2	ADAE	USUBJID	受试者唯一标识符	Unique Subject Identifier			
4	3	ADAE	SUBJID	受试者标识符	Subject Identifier for the Study			
5	4	ADAE	SITEID	研究中心标识符	Study Site Identifier			
6	5	ADAE	SITEIDSS	特定研究中心标识符	Study Specific Site Identifier			
7	6	ADAE	COUNTRYL	国家/地区（全称）	Country (Full Name)			
23	22	ADAE	TRTPN	计划治疗（N）	Planned Treatment (N)			
25	24	ADAE	TRTAN	实际治疗（N）	Actual Treatment (N)			
26	25	ADAE	AESPID	申办方定义的标识符	Sponsor-Defined Identifier			
28	27	ADAE	AECAT	不良事件类别	Category for Adverse Event			
29	28	ADAE	AETERM	不良事件的报告名称	Reported Term for the Adverse Event			
42	41	ADAE	AESTDTC	不良事件的开始日期/时间	Start Date/Time of Adverse Event			
43	42	ADAE	AESTDY	不良事件开始的研究日	Study Day of Start of Adverse Event			
46	45	ADAE	AEENDTC	不良事件的结束日期/时间	End Date/Time of Adverse Event			
47	46	ADAE	AEENDY	不良事件结束的研究日	Study Day of End of Adverse Event			
52	51	ADAE	ADURN	持续时间（N）-分析用	Analysis Duration (N)			
58	57	ADAE	AETOXGR	标准毒性分级	Standard Toxicity Grade			
61	60	ADAE	AEACNOTH	采取的其他措施	Other Action Taken			
62	61	ADAE	AESCONG	导致先天性异常或出生缺陷	Congenital Anomaly or Birth Defect			
63	62	ADAE	AESDISAB	导致永久或显著的残疾/失能	Persist or Signify Disability/Incapacity			
65	64	ADAE	AESMIE	其他具有重要医学意义的严重事件	Other Medically Important Serious Event			
66	65	ADAE	AEOUT	不良事件的结局	Outcome of Adverse Event			
85	84	ADAE	ATRTSDTC	首次治疗的日期/时间-分析用	Analysis Date/Time of First Exp to Trt.			
87	86	ADAE	ATRTSDTF	首次治疗填补标记的日期-分析用	Analysis Date of First Exp Imput. Flag			

Table 1 Partial Records From ADAE provided by translator vendor Before CMAT Process

	A	B	C	D	E	W	X	Y
1	Order	Dataset	Variable	Label in Chinese	Label	CSP_Comments	Vendor_Text	CSP_Discussion_YN
2	1	ADAE	STUDYID	研究标识符	Study Identifier			
3	2	ADAE	USUBJID	受试者唯一标识符	Unique Subject Identifier			
4	3	ADAE	SUBJID	受试者标识符	Subject Identifier for the Study			
5	4	ADAE	SITEID	研究中心标识符	Study Site Identifier			
6	5	ADAE	SITEIDSS	特定研究中心标识符	Study Specific Site Identifier			
7	6	ADAE	COUNTRYL	国家/地区（全称）	Country (Full Name)			
23	22	ADAE	TRTPN	计划治疗(N)	Planned Treatment (N)	Follow Meta Translation File	计划治疗（N）	
25	24	ADAE	TRTAN	实际治疗(N)	Actual Treatment (N)	Follow Meta Translation File	实际治疗（N）	
26	25	ADAE	AESPID	自定义标识符	Sponsor-Defined Identifier	Follow Meta Translation File	申办方定义的标识符	
28	27	ADAE	AECAT	类别	Category for Adverse Event	Follow Meta Translation File	不良事件类别	
29	28	ADAE	AETERM	不良事件报告名称	Reported Term for the Adverse Event	Follow Meta Translation File	不良事件的报告名称	
42	41	ADAE	AESTDTC	开始日期/时间	Start Date/Time of Adverse Event	Follow Meta Translation File	不良事件的开始日期/时间	
43	42	ADAE	AESTDY	开始日	Study Day of Start of Adverse Event	Follow Meta Translation File	不良事件开始的研究日	
46	45	ADAE	AEENDTC	结束日期/时间	End Date/Time of Adverse Event	Follow Meta Translation File	不良事件的结束日期/时间	
47	46	ADAE	AEENDY	结束日	Study Day of End of Adverse Event	Follow Meta Translation File	不良事件结束的研究日	
52	51	ADAE	ADURN	持续时间(N)-分析用	Analysis Duration (N)	Follow Meta Translation File	持续时间（N）-分析用	
58	57	ADAE	AETOXGR	毒性分级	Standard Toxicity Grade	Follow Meta Translation File	标准毒性分级	
61	60	ADAE	AEACNOTH	采取的其它措施	Other Action Taken	Follow Meta Translation File	采取的其他措施	
62	61	ADAE	AESCONG	先天性异常或出生缺陷	Congenital Anomaly or Birth Defect	Follow Meta Translation File	导致先天性异常或出生缺陷	
63	62	ADAE	AESDISAB	永久或严重的残疾或功能丧失	Persist or Signify Disability/Incapacity	Follow Meta Translation File	导致永久或显著的残疾/失能	
65	64	ADAE	AESMIE	其他重要的医学事件	Other Medically Important Serious Event	Follow Meta Translation File	其他具有重要医学意义的严重事件	
66	65	ADAE	AEOUT	转归	Outcome of Adverse Event	Follow Meta Translation File	不良事件的结局	
85	84	ADAE	ATRTSDTC	首次治疗的日期/时间-分析用	Analysis Date/Time of First Exp to Trt.	Meeting length<=40 rule	首次治疗的日期/时间-分析用	
87	86	ADAE	ATRTSDTF	首次治疗填补标记的日期-分析用	Analysis Date of First Exp Imput. Flag	Meeting length<=40 rule	首次治疗填补标记的日期-分析用	

Table 2 Partial Records From ADAE provided by translator vendor After CMAT Process

For EXCEL SPEC sheets including 'ValueLevel', 'Methods', 'Comments' and other related files, the CSP shall conduct a comprehensive review to ensure terminological consistency across all documents for identical items. Additionally, the CSP will facilitate the conversion of aCRF, CSDRG, and ADRG into PDF format with bookmarks and configure file properties as required. The CSP must also ensure that the Chinese version of aCRF maintains one-to-one page correspondence with the English version; failure to do so will necessitate updates to the Chinese define.xml files for all affected page references.

Chinese Package Labels Replacement

Upon availability of the quality-reviewed EXCEL SPEC, the CSP shall update the labels for datasets, variables, and QLABEL in SUPPxx domains and verify that special characters are compatible with UTF-8 encoding. It is strongly recommended to conduct a comparative analysis of the Chinese datasets against the original English datasets to confirm that only expected differences are observed, followed by generation of XPT files for all Chinese datasets.

Chinese Package Finalization

The CSP shall import the Chinese aCRF, XPT files, and EXCEL SPEC into P21E to generate the Chinese version of define.xml files and P21E reports. The CSP will then address P21E issues triggered by the NMPA validation engine and export the define.xml files for subsequent review and integration. For recurring issues identified across different P21E validation engines, the CSP may resolve these messages using Domain and Rule ID as keys from the issue summary sections of the translated review guides. For new issues, comprehensive investigation and documentation will be required. The CSP shall integrate and finalize CSDRG and ADRG with all issues identified by the NMPA validation engine and update the issue summary accordingly. Additionally, the CSP will update CSDRG Section 4.1 and ADRG Section 6.1 to document the use of the NMPA validation engine and subsequently archive and close the analysis folder in accordance with study conventions.

CONCLUSION

The preparation of NMPA submission packages encompasses multiple procedural steps and responsibilities within the global pharmaceutical environment. Comprehensive quality review and proofreading processes enhance the quality, reliability, and readability of Chinese submission packages. Furthermore, the CMAT tool facilitates the quality review process and enables efficient standardization of dataset and variable labels across diverse product teams. Effective team collaboration is essential to ensure successful completion of the NMPA submission package preparation process.

REFERENCES

1 NMPA, 20200720. 国家药监局药审中心关于发布《药物临床试验数据递交指导原则（试行）》的通告（2020 年第 16 号）
https://www.cde.org.cn/main/news/viewInfoCommon/f649995d3a9ade8dcd67f6a2ced36f0b

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Fu Wang
Parexel International
Fu.Wang@parexel.com